CONTENTS

1 Working with Matrices and Vectors

1.1 Notation for Scalars, Vectors, and Matrices

Lowercase letters $\Rightarrow$ scalars: x; c; σ.

Boldface, lowercase letters $\Rightarrow$ vectors: $\mathbf{x}$; $\mathbf{y}$; $\boldsymbol{\beta}$.

Boldface, uppercase letters $\Rightarrow$ matrices: $\mathbf{A}$; $\mathbf{X}$; $\boldsymbol{\Sigma}$.

1.2 Matrix and Vector Operations

Definition 1.
A column of real numbers is called a **vector**.

Example 1.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{a} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Since $\mathbf{y}$ has n elements it is said to have **order** (or dimension) n.

Definition 2.
A rectangular array of elements with m rows and k columns is called an $m \times k$ **matrix.**

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix}$$

This matrix is said to be of **order** (or dimension) $m \times k$, where

- m is the **row** order (dimension)
- k is the **column** order (dimension)

Example 2.

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & -2 \\ 0 & 4 & 5 \end{bmatrix} \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

```
R-code:
A = matrix(c(1,3,-2,0,4,5), 2,3, byrow  = T)
A
I = diag(rep(1,3))
I
B = matrix(c(1,3,2,6), 2,2, byrow  = T)
B
```

Definition 3. Matrix addition
If $\mathbf{A}$ and $\mathbf{B}$ are both $m \times k$ matrices, then

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$
$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mk} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1k}+b_{1k} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2k}+b_{2k} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mk}+b_{mk} \end{bmatrix}$$

Notation:

$$C_{m\times k} = \{c_{ij}\} \text{ where } c_{ij} = a_{ij} + b_{ij}$$

Definition 4. Matrix subtraction

If $\mathbf{A}$ and $\mathbf{B}$ are $m \times k$ matrices, then $\mathbf{C} = \mathbf{A} - \mathbf{B}$ is defined by

$$\mathbf{C} = \{c_{ij}\} \text{ where } c_{ij} = a_{ij} - b_{ij}\ .$$

Example 3.

$$\begin{bmatrix} 3 & 6 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 7 & -4 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ -1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}$$

R-codes:

```
A = matrix(c(3,6,2,1), 2,2,byrow=T)
B = matrix(c(7,-4,-3,2), 2,2,byrow=T)
C = A+B
D = matrix(c(1,-1,1,1,1,0), 3,2,byrow=T)
E = matrix(c(1,-1,2,0,1,1), 3,2,byrow=T)
F = D-E
```

Definition 5. Scalar multiplication

Let a be a scalar and $\mathbf{B} = \{b_{ij}\}$ be an $m \times k$ matrix, then

$$a\,\mathbf{B} = \mathbf{B}\,a = \{a\,b_{ij}\}$$

Example 4.

$$2\begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 4 & -2 & 6 \\ 0 & 8 & -4 \end{bmatrix}$$

```
 R-Code:
A = matrix(c(2,-1,3,0,4,2), 2,3,byrow=T)
B= 2*A
```


Definition 6. Transpose
The transpose of the $m \times k$ matrix $\mathbf{A} = \{a_{ij}\}$ is the $k \times m$ matrix with elements $\{a_{ji}\}$. The transpose of $\mathbf{A}$ is denoted by $\mathbf{A^T}$ (or $\mathbf{A}'$).

Example 5.

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 3 & 0 \\ -2 & 6 \end{bmatrix} \qquad \mathbf{A^T} = \begin{bmatrix} 1 & 3 & -2 \\ 4 & 0 & 6 \end{bmatrix}$$

```
R-code:
A = matrix(c(1,4,3,0,-2,6), 3,2,byrow=T)
AT = t(A)
AT
```

Definition 7. If a matrix has the same number of rows and columns it is called a **square matrix**.

$$\mathbf{A}_{k\times k} = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix}$$

is said to have order (or dimension) k.

Definition 8. A square matrix $\mathbf{A} = \{a_{ij}\}$ is **symmetric** if $\mathbf{A} = \mathbf{A}^T$, that is, if $a_{ij} = a_{ji}$ for all (i, j).

Example 6.

$$\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 4 & 3 & 2 & 1 \\ 3 & 5 & 0 & -2 \\ 2 & 0 & 3 & -1 \\ 1 & -2 & -1 & 2 \end{bmatrix}$$

Definition 9. Inner product (crossproduct) of two vectors of order n

$$\begin{aligned}\mathbf{a}^T\mathbf{y} &= [a_1, a_2, \cdots a_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= [a_1y_1 + a_2y_2 + \cdots + a_ny_n \\ &= \sum_{j=1}^{n} a_jy_j\end{aligned}$$

Note that $\mathbf{a}^T\mathbf{y} = \mathbf{y}^T\mathbf{a}$

```
R-codes:
a = c(1, 7, -6, 4)
y = c(2,-2,1,5)
aTy1 = t(a)%*%y
aTy2 = a%*%y
aTy3 = crossprod(a,y)
```

Definition 10. Euclidean distance (or length of a vector)

$$\|\mathbf{y}\| = (\mathbf{y}^T\,\mathbf{y})^{1/2} = \left(\sum_{j=1}^{n} y_j^2\right)^{1/2}$$

```
R-Code:
y = c(2,-2,1,5)
ynorm = sqrt(crossprod(y,y))
ynorm
```


Definition 11. Matrix multiplication
The product of an $n \times k$ matrix $\mathbf{A}$ and a $k \times m$ matrix $\mathbf{B}$ is the $n \times m$ matrix $\mathbf{C} = \{c_{ij}\}$ with elements

$$c_{ij} = a_{i1}\,b_{1j} + a_{i2}\,b_{2j} + \cdots + a_{ik}\,b_{kj}$$

Example 7.

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & -2 \\ 1 & -1 & 4 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \quad \mathbf{C} = \mathbf{AB} = \begin{bmatrix} 1 & -3 \\ 4 & 11 \end{bmatrix}$$

```
R-codes:
A = matrix(c(3,0,-2,1,-1,4), 2,3,byrow=T)
B = matrix(c(1,1,1,2,1,3), 3,2,byrow=T)
C = A%*%B
```

Definition 12. Elementwise multiplication of two matrices

$$\mathbf{A} \# \mathbf{B} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix} \# \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{km} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}\, b_{11} & \cdots & a_{1m}\, b_{1m} \\ \vdots & & \vdots \\ a_{k1}\, b_{k1} & \cdots & a_{km}\, b_{km} \end{bmatrix}$$

Example 8.

$$\begin{bmatrix} 3 & 1 \\ 2 & 4 \\ 0 & 6 \end{bmatrix} \# \begin{bmatrix} 1 & -5 \\ -3 & 4 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ -6 & 16 \\ 0 & 12 \end{bmatrix}$$

```
R-codes:
A = matrix(c(3,1,2,4,0,6), 3,2,byrow=T)
B = matrix(c(1,-5,-3,4,-2,2), 3,2,byrow=T)
C = A*B
```


Definition 13. Kronecker product of two matrices

$$\mathbf{A}_{k\times m} \otimes \mathbf{B}_{n\times s} = \begin{bmatrix} a_{11}\,\mathbf{B} & a_{12}\,\mathbf{B} & \cdots & a_{1m}\,\mathbf{B} \\ a_{21}\,\mathbf{B} & a_{22}\,\mathbf{B} & \cdots & a_{2m}\,\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{k1}\,\mathbf{B} & a_{k2}\,\mathbf{B} & \cdots & a_{km}\,\mathbf{B} \end{bmatrix}$$

Example 9.

$$\begin{bmatrix} 2 & 4 \\ 0 & -2 \\ 3 & -1 \end{bmatrix} \otimes \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 20 & 12 \\ 4 & 2 & 8 & 4 \\ 0 & 0 & -10 & -6 \\ 0 & 0 & -4 & -2 \\ 15 & 9 & -5 & -3 \\ 6 & 3 & -2 & -1 \end{bmatrix}$$

$$\mathbf{a} \otimes \mathbf{y} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \otimes \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_1 y_1 \\ a_1 y_2 \\ a_2 y_1 \\ a_2 y_2 \\ a_3 y_1 \\ a_3 y_2 \end{bmatrix}$$

```
R-codes:
A = matrix(c(2,4,0,-2,3,-1),ncol=2,byrow=T)
B = matrix(c(5,3,2,1),2,2,byrow=T)
C = kronecker(A,B)
```


1.3 Determinant

Definition 14. The **determinant** of an $n \times n$ matrix $\mathbf{A}$ is

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}| \quad \text{for any row } i$$

or

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|M_{ij}| \quad \text{for any column } j$$

where M_{ij} is the "minor" for a_{ij} obtained by deleting the i^{th} row and j^{th} column from A.

Example 10.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$|\mathbf{A}| = a_{11}(-1)^{1+1}|a_{22}| + a_{12}(-1)^{1+2}|a_{21}|$$

then $\begin{vmatrix} 7 & 2 \\ 4 & 5 \end{vmatrix} =$

Example 11.

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$
$$= a_{11}(-1)^{1+1}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^{1+2}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(-1)^{1+3}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

then $\begin{vmatrix} 1 & 1 & 3 \\ 4 & 3 & 6 \\ 7 & 5 & 9 \end{vmatrix} =$

```
R-codes:
A = matrix(c(1,1,3,4,3,6,7,5c‘,9),3,3,byrow=T)
> detA = det(A)
> detA
```


Properties of determinants

- $|\mathbf{A}^{\mathbf{T}}| = |\mathbf{A}|$
- $|\mathbf{A}|$ = product of the eigenvalues of $\mathbf{A}$
- $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ when $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same order.
- $\begin{vmatrix} \mathbf{P} & 0 \\ \mathbf{X} & \mathbf{Q} \end{vmatrix} = |\mathbf{P}||\mathbf{Q}|$ when $\mathbf{P}$ and $\mathbf{Q}$ are square matrices of the same order and 0 is a matrix of zeros.
- $|\mathbf{AB}| = |\mathbf{BA}|$ when the matrix product is defined
- $|c\mathbf{A}| = c^k|\mathbf{A}|$ when c is a scalar and $\mathbf{A}$ is a $k \times k$ matrix

1.4 Orthogonal and Idempotent Matrices

Definition 15. A square matrix $\mathbf{A}$ is said to be **orthogonal** if

$$\mathbf{A}\mathbf{A}^{\mathbf{T}} = \mathbf{A}^{\mathbf{T}}\mathbf{A} = I \quad (\text{then } \mathbf{A}^{-1} = \mathbf{A}^{\mathbf{T}})$$

Definition 16. A square matrix $\mathbf{P}$ is **idempotent** if $\mathbf{PP} = \mathbf{P}$

Example 12. (Orthogonal Matrix)

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

In each case the columns of $\mathbf{A}$ are coefficients for orthogonal contrasts.

Example 13. (Idempotent Matrix)

$$\mathbf{P} = \begin{bmatrix} \frac{5}{6} & \frac{2}{6} & -\frac{1}{6} \\ \frac{2}{6} & \frac{2}{6} & \frac{2}{6} \\ -\frac{1}{6} & \frac{2}{6} & \frac{5}{6} \end{bmatrix}$$

Example 14. Use the definition of orthogonal and idempotent matrices and properties of determinants to prove the following results:

(a) If $\mathbf{A}$ is an orthogonal matrix, then $|\mathbf{A}|$ is either 1 or -1. (Hint: use the definition of an orthogonal matrix and consider the determinant of an identity matrix.)

(b) If $\mathbf{W}$ is an idempotent matrix, then $|\mathbf{W}|$ is either 0 or 1.

1.5 Linear Combinations and Column Spaces

$\mathbf{Ab}$ is a linear combination of the columns of an $m \times n$ of matrix $\mathbf{A}$.

$$\mathbf{Ab} = \begin{bmatrix}\mathbf{a_1}, \ldots, \mathbf{a_n}\end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = b_1\mathbf{a_1} + \cdots + b_n\mathbf{a_n}$$

The set of all possible linear combinations of the columns of $\mathbf{A}$ is called the column space of $\mathbf{A}$ and is written as

$$\mathcal{C}(\mathbf{A}) = \{\mathbf{Ab} : \mathbf{b} \in \mathbf{R}^n\}$$

Note that $\mathcal{C}(\mathbf{A}) \subseteq \mathbf{R}^m$.

Example 15. Suppose $\boldsymbol{X}$ is an $n \times p$ matrix and $\boldsymbol{B}$ is a $p \times p$ non-singular matrix. Prove that $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$.

1.6 Linear Independence

Definition 17. A set of n-dimensional vectors $\mathbf{y}_1, \mathbf{y}_2, \cdots \mathbf{y}_k$ are **linearly independent** if there is no set of scalars $a_1\ a_2\ \cdots\ a_k$ such that

$$\mathbf{0} = \sum_{j=1}^{k} a_j\, \mathbf{y}_j$$

and at least one a_j is non-zero.

Example 16. Show that

$$\mathbf{y}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{y}_2 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad \mathbf{y}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

are linearly independent.

Example 17. Show that

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \mathbf{y}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{y}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

are not linearly independent.

1.7 Rank

Definition 18. The **row rank** of a matrix is the number of linearly independent rows, where each row is considered as a vector.

Definition 19. The **column rank** of a matrix is the number of linearly independent columns, with each column considered as a vector.

Example 18. Show that the row and column rank of

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

is 2.

```
R-codes:
A = matrix(c(1,1, 1,2,5,-1,0,1,-1),3,3,byrow=T)
rA = qr(A)$rank
```


Result 1. The row rank and the column rank of a matrix are equal.

Definition 20. The **rank** of a matrix is either the row rank or the column rank of the matrix.

Definition 21. A square matrix $\mathbf{A}_{k\times k}$ is **nonsingular** if its rank is equal to the number of rows (or columns).
This is equivalent to the condition

$$\mathbf{A}_{k\times k}\,\mathbf{b}_{k\times 1} = \mathbf{0}_{k\times 1} \text{ only when } \mathbf{b} = \mathbf{0}$$

A matrix that fails to be nonsingular is called **singular**.

Result 2. If $\mathbf{B}_{n\times n}$ is non-singular and $\mathbf{A}_{n\times m}$, then

$$rank(\mathbf{BA}) = rank(\mathbf{A}).$$

Result 3. If $\mathbf{B}$ and $\mathbf{C}$ are non-singular matrices and products with $\mathbf{A}$ are defined, then

$$rank(\mathbf{BA}) = rank(\mathbf{AC}) = rank(\mathbf{A}).$$

Result 4. $rank(\mathbf{A}^T\mathbf{A}) = rank(\mathbf{A}) = rank(\mathbf{A}^T)$.

1.8 Inverse

Definition 22. The **identity matrix**, denoted by $\mathbf{I}$, is a $k \times k$ matrix of the form

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Definition 23. The **inverse** of a square, non-singular matrix $\mathbf{A}$ is the matrix, denoted by $\mathbf{A}^{-1}$, such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = I$$

```
R-codes:
I3 = diag(rep(1,3))
I3
W = matrix(c(1,2,3,4,5,6,7,8,10),3,3,byrow=T)
Winv = solve(W)
Winv
```

Result 5.

(i) The inverse of $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

(ii) In general, the (i, j) element of $\mathbf{A}^{-1}$ is

$$\frac{(-1)^{i+j}\ |\mathbf{A}_{ji}|}{|\mathbf{A}|}$$

where $\mathbf{A}_{ji}$ is the matrix obtained by deleting the j-th row and i-th column of $\mathbf{A}$.

Result 6. For a $k \times k$ matrix $\mathbf{A}$, the following are equivalent:

(i) $\mathbf{A}$ is nonsingular

(ii) $|\mathbf{A}| \neq 0$

(iii) $\mathbf{A}^{-1}$ exists

Result 7. For $k \times k$ nonsingular matrices $\mathbf{A}$ and $\mathbf{B}$

(i) $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$

(ii) $(\mathbf{A}\,\mathbf{B})^{-1} = \mathbf{B}^{-1}\,\mathbf{A}^{-1}$

(iii) $|\mathbf{A}^{-1}| = 1/|\mathbf{A}|$

(iv) $\mathbf{A}^{-1}$ is unique and nonsingular

(v) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$

(vi) If $\mathbf{A}$ is symmetric, than $\mathbf{A}^{-1}$ is symmetric

Example 19.

Suppose $\boldsymbol{X}$ is an $n \times p$ matrix and $\boldsymbol{B}$ is a $p \times p$ non-singular matrix. Prove that $\mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}\boldsymbol{B}^{-1})$.

Result 8. Inverse of a Diagonal Matrix

$$\begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{kk} \end{bmatrix}^{-1} = \begin{bmatrix} 1/a_{11} & & & \\ & 1/a_{22} & & \\ & & \ddots & \\ & & & 1/a_{kk} \end{bmatrix}$$

Result 9.

If $\mathbf{B}$ is a $k \times k$ non-singular matrix and $\mathbf{B} + \mathbf{c}\mathbf{c}^T$ is non-singular, then

$$(\mathbf{B} + \mathbf{c}\mathbf{c}^T)^{-1} = \mathbf{B}^{-1} - \frac{\mathbf{B}^{-1}\mathbf{c}\mathbf{c}^T\mathbf{B}^{-1}}{1 + \mathbf{c}^T\mathbf{B}^{-1}\mathbf{c}}$$

Result 10.

Let $\mathbf{I}_n$ be an $n \times n$ identity matrix and let $\mathbf{J}_n = \mathbf{1}\mathbf{1}^T$ be an $n \times n$ matrix where each element is one, then

$$(a\,\mathbf{I}_n + b\,\mathbf{J}_n)^{-1} = \frac{1}{a}\left(\mathbf{I}_n - \frac{b}{a+nb}\mathbf{J}_n\right)$$

Example 20.

Suppose $\mathbf{Z} = \underset{3\times1}{\mathbf{1}}$, $\mathbf{G} = 9\underset{3\times3}{\mathbf{I}}$, $\mathbf{R} = 25\underset{3\times3}{\mathbf{I}}$. If $\mathbf{\Sigma} = \mathbf{ZGZ^T} + \mathbf{R}$, find $\mathbf{\Sigma^{-1}}$.

1.9 Trace

Definition 24. The **trace** of a $k \times k$ matrix $\mathbf{A} = \{a_{ij}\}$ is the sum of the diagonal elements:

$$tr(\mathbf{A}) = \sum_{j=1}^{k} a_{jj}$$

```
R-codes:
W = matrix(c(1,2,3,4,5,6,7,8,10),
            3,3,byrow=T)
trW = sum(diag(W))
trW
```

```
R-codes:
W = {1 2 3, 4 5 6, 7 8 10};
trW1 = trace(W);
trW2 = sum(diag(W));
print W trW1 trW2;
```

Result 11. Let $\mathbf{A}$ and $\mathbf{B}$ denote $k \times k$ matrices and let c be a scalar. Then,

(i) $tr(c\,\mathbf{A}) = c\,tr(\mathbf{A})$
(ii) $tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$
(iii) $tr(\mathbf{AB}) = tr(\mathbf{BA})$
(iv) $tr(\mathbf{B}^{-1}\mathbf{A}\,\mathbf{B}) = tr(\mathbf{A})$
(v) $tr(\mathbf{A}\,\mathbf{A}^T) = \displaystyle\sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}^2$

Example 21.
For $\mathbf{A} = \mathbf{I}_{n\times n} - \frac{1}{n}\mathbf{i}\mathbf{i}^{\mathbf{T}}$ where $\mathbf{i}^{\mathbf{T}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}_{1\times n}$.

(a) Show that $\mathbf{A}$ is idempotent.

(b) Find $tr(\mathbf{A})$.

(c) Interpret the result of $\mathbf{Ay}$ where $\mathbf{y}$ is $n \times 1$.

1.10 Eigenvalues and Eigenvectors

Definition 25. For a $k \times k$ matrix $\mathbf{A}$, the scalars $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ satisfying the polynomial equation

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

are called the eigenvalues (or characteristic roots) of $\mathbf{A}$.

Definition 26. Corresponding to any eigenvalue λ_i is an eigenvector (or characteristic vector) $\mathbf{u}_i \neq \mathbf{0}$ satisfying

$$\mathbf{A}\,\mathbf{u}_i = \lambda_i\,\mathbf{u}_i.$$

Comment: Eigenvectors are not unique

(i) If $\mathbf{u}_i$ is an eigenvector for λ_i, then $c\,\mathbf{u}_i$ is also an eigenvector for any scalar $c \neq 0$.

(ii) We will adopt the following conventions (for real symmetric matrices)

$$\mathbf{u}_i^T\mathbf{u}_i = 1 \qquad \text{for all} i = 1, \ldots, k$$

$$\mathbf{u}_i^T\mathbf{u}_j = 0 \qquad \text{for all} i \neq j$$

(iii) Even with (ii), eigenvectors are not unique

- If $\mathbf{u}_i$ is an eigenvector satisfying (ii), then $-\mathbf{u}_i$ is also an eigenvector satisfying (ii).
- If $\lambda_i = \lambda_j$ then there are an infinite number of choices for $\mathbf{u}_i$ and $\mathbf{u}_j$.

Example 22. Find the eigenvalue and eigenvector of

$$\mathbf{A} = \begin{bmatrix} 1.96 & 0.72 \\ 0.72 & 1.54 \end{bmatrix}$$

using R.

Result 12. For a $k \times k$ symmetric matrix $\mathbf{A}$ with elements that are real numbers

(i) every eigenvalue of $\mathbf{A}$ is a real number

(ii) rank($\mathbf{A}$) = number of non-zero eigenvalues

(iii) if $\mathbf{A}$ is non-negative definite, then $\lambda_i \geq 0$ for all $i = 1, 2, \ldots, k$

(iv) if $\mathbf{A}$ is positive definite then $\lambda_i > 0$ for all $i = 1, 2, \ldots, k$

(v) $\text{trace}(\mathbf{A}) = \sum_{i=1}^{k} a_{ii} = \sum_{i=1}^{k} \lambda_i$

(vi) $|\mathbf{A}| = \Pi_{i=1}^{k} \lambda_i$

(vii) if $\mathbf{A}$ is idempotent ($\mathbf{A}\,\mathbf{A} = \mathbf{A}$), then the eigenvalues are either zero or one.

Example 23.
Find the rank of the idempotent matrix $X(X^TX)^{-1}X^T$ where X is $n \times p$ and X^TX is nonsingular.

1.11 Quadratic Form

Definition 27.
Let $\mathbf{A}$ be a $k \times k$ matrix and let $\mathbf{y}$ be a vector of order k, then

$$\mathbf{y}^T\mathbf{A}\mathbf{y} = \sum_{i=1}^{k}\sum_{j=1}^{k} y_i y_j a_{ij}$$

is called a **quadratic form**.

Suppose $\mathbf{y}_{n\times 1}$ is a vector of n observations. Then $\mathbf{y}'\mathbf{y} = \sum_{i=1}^{n} y_i^2$ is the total sum of squares of the observations. Let $\mathbf{P}$ be an orthogonal matrix

$$\mathbf{P}\mathbf{P}' = \mathbf{P}'\mathbf{P} = \mathbf{I}$$

and partition $\mathbf{P}$ row wise into k sub-matrices $\mathbf{P}_i$, of order $n_i \times n$, for $i = 1, 2, \ldots, k$, with $\sum_{i=1}^{k} n_i = n$; i.e.

$$\mathbf{P} \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \vdots \\ \mathbf{P}_k \end{bmatrix} \quad \text{and} \quad \mathbf{P}' = \begin{bmatrix} \mathbf{P}_1' & \mathbf{P}_2' & \cdots & \mathbf{P}_k' \end{bmatrix}.$$

Then $\mathbf{y}'\mathbf{y} = \mathbf{y}'\mathbf{I}\mathbf{y} = \mathbf{y}'\mathbf{P}'\mathbf{P}\mathbf{y} = \sum_{i=1}^{k} \mathbf{y}'\mathbf{P}_i'\mathbf{P}_i\mathbf{y}$.

In this way $\mathbf{y}'\mathbf{y}$ is partition into k sums of squares

$$\mathbf{y}'\mathbf{P}_i'\mathbf{P}_i\mathbf{y} \quad \text{for} \quad i = 1, \ldots, k$$

each of these sums of squares corresponds to the lines in an analysis of variance, having $\mathbf{y}'\mathbf{y}$ as the total sums of squares.

Example 24.
Corresponding to a vector of 4 observations consider

$$\mathbf{P} = \left[\begin{array}{cccc} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} \\ \hline \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{-3}{\sqrt{12}} \end{array}\right] = \left[\begin{array}{c} \mathbf{P}_1 \\ \hline \mathbf{P}_2 \end{array}\right]$$

Show that $\mathbf{P}$ is orthogonal and find the two partition sums of squares.

1.11.1 Symmetric Matrices

Any quadratic form $\mathbf{y}^T\mathbf{A}\mathbf{y}$ can be written as $\mathbf{y}^T\mathbf{A}\mathbf{y} = \mathbf{y}^T\mathbf{B}\mathbf{y}$ where $\mathbf{B} = \frac{1}{2}(\mathbf{A}+\mathbf{A}^T)$ is symmetric. Furthermore, any quadratic form can be written as $\mathbf{y}^T\mathbf{A}\mathbf{y}$ for an infinite number of matrices, but can only be written in one way as $\mathbf{y}^T\mathbf{B}\mathbf{y}$ for $\mathbf{B}$ symmetric. For example,

$$4y_1^2+6y_1y_2+7y_1^2 = \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 4 & 3+a \\ 3-a & 7 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

for any value of a, but only when $a = 0$ is the matrix involved symmetric. This means that for any particular quadratic form there is only one, unique matrix such that the quadratic form can be written as $\mathbf{y}^T\mathbf{A}\mathbf{y}$ with $\mathbf{A}$ being symmetric. Due to the uniqueness of this symmetric matrix, the quadratic form that we are going to discuss is confined to the case of $\mathbf{A}$ being symmetric.

1.11.2 Positive Definiteness

Definition 28.
A quadratic form $\mathbf{y}^T\mathbf{A}\mathbf{y}$ is said to be **positive definite** (p.d.)if

$$\mathbf{y}^T\mathbf{A}\mathbf{y} > 0 \text{ for all } \mathbf{y} \text{ except} \mathbf{y} = \mathbf{0}.$$

The corresponding (symmetric) matrix is also described as positive definite.

Definition 29. A quadratic form $\mathbf{y}^T\mathbf{A}\mathbf{y}$ is said to be **positive semi-definite** (p.s.d) if

$$\mathbf{y}^T\mathbf{A}\mathbf{y} \geq 0 \quad \text{for all } \mathbf{y} \neq \mathbf{0}$$
$$\text{with} \mathbf{y}^T\mathbf{A}\mathbf{y} = 0 \text{ for at least one} \mathbf{y} \neq \mathbf{0}.$$

The corresponding (symmetric) matrix $\mathbf{A}$ is a p.s.d. matrix.

Example 25. Show that

$$\mathbf{A} = \begin{pmatrix} 3 & 5 & 1 \\ 5 & 13 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

is a positive definite matrix.

Example 26.
Show that

$$\mathbf{B} = \begin{pmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix}$$

is a positive semidefinite matrix.

Example 27. Let $\mathbf{B}$ be an $n \times p$ matrix. Show that

(i) If $rank(\mathbf{B}) = p$, then $\mathbf{B}^T\mathbf{B}$ is positive definite.

(ii) If $rank(\mathbf{B}) < p$, then $\mathbf{B}^T\mathbf{B}$ is positive semidefinite.

1.12 Spectral Decomposition

Result 13. The spectral decomposition of a $k \times k$ symmetric matrix $\mathbf{A}$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ and eigenvectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ (with $\mathbf{u}_i^T\,\mathbf{u}_i = 1$ and $\mathbf{u}_i^T\,\mathbf{u}_j = 0$) is

$$\begin{aligned}\mathbf{A} &= \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2\,\mathbf{u}_2\,\mathbf{u}_2^T + \cdots + \lambda_k\,\mathbf{u}_k\,\mathbf{u}_k^T \\ &= \mathbf{U}\,\mathbf{D}\,\mathbf{U}^T\end{aligned}$$

where

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_k \end{bmatrix}$$

and

$$\mathbf{U} = [\mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_k\,]$$

is an orthogonal matrix.

Result 14. If $\mathbf{A}$ is a $k \times k$ symmetric nonsingular matrix with spectral decomposition

$$\mathbf{A} = \sum_{i=1}^{k} \lambda_i \, \mathbf{u}_i \, \mathbf{u}_i^T = \mathbf{U} \, \mathbf{D} \, \mathbf{U}^T$$

then

(i) $\mathbf{A}^{-1} = \sum_{i=1}^{k} \lambda_i^{-1} \, \mathbf{u}_i \, \mathbf{u}_i^T = \mathbf{U} \, \mathbf{D}^{-1} \, \mathbf{U}^T$

(ii) the square root matrix

$$\mathbf{A}^{1/2} = \sum_{i=1}^{k} \sqrt{\lambda_i} \, \mathbf{u}_i \, \mathbf{u}_i^T$$

has the properties:

(a) $\mathbf{A}^{1/2} \, \mathbf{A}^{1/2} = \mathbf{A}$

(b) $\mathbf{A}^{1/2} \, \mathbf{A}^{-1} \, \mathbf{A}^{1/2} = I$

(c) $\mathbf{A}^{1/2}$ is symmetric

(iii) The inverse square root matrix

$$\begin{aligned} \mathbf{A}^{-1/2} &= \sum_{i=1}^{k} \frac{1}{\sqrt{\lambda_i}} \mathbf{u}_i \, \mathbf{u}_i^T \\ &= \mathbf{U} \mathbf{D}^{-1/2} \mathbf{U}^T \end{aligned}$$

has the properties:

(a) $\mathbf{A}^{-1/2} \, \mathbf{A}^{-1/2} = \mathbf{A}^{-1}$

(b) $\mathbf{A}^{-1/2} \, \mathbf{A} \, \mathbf{A}^{-1/2} = I$

(c) $\mathbf{A}^{-1/2}$ is symmetric

In parts (ii) and (iii), $\mathbf{A}$ should be positive definite to ensure that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$$

Example 28.
Let $\mathbf{A}$ be a Positive definite matrix, show that there exists a nonsingular $\mathbf{H}$ such that $\mathbf{A} = \mathbf{H}^2$.

1.13 Random Vectors:

Definition 30.

A random vector $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ is a vector whose elements are random variables.

1.13.1 Mean vectors:

$$E(\mathbf{y}) = \begin{bmatrix} E(y_1) \\ \vdots \\ E(y_n) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = \boldsymbol{\mu}$$

where

$$\mu_i = E(y_i) = \int_{-\infty}^{\infty} y f_i(y) dy$$

if y_i is a continuous random variable with density function $f_i(y)$
and

$$\mu_i = E(y_i) = \sum y p_i(y)$$

if y_i is a discrete random variable with probability function $p_i(y)$.

1.13.2 Covariance matrix:

$$\mathbf{\Sigma} = Var(\mathbf{y}) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_n^2 \end{bmatrix}$$

with variances

$$Var(y_i) = \sigma_i^2 = E(y_i - \mu_i)^2$$

$$= \begin{bmatrix} \int_{-\infty}^{\infty} (y - \mu_i)^2 f_i(y) dy & \text{if } y \text{ is a continuous} \\ & \text{random variable} \\ \sum_{all\ y} (y - \mu_i)^2 p_i(y) & \text{if } y \text{ is a discrete} \\ & \text{random variable} \end{bmatrix}$$

and covariances:

$$\sigma_{ij} = Cov(y_i, y_j) = E\left[(y_i - \mu_i)(y_j - \mu_j)\right]$$

where

$$\sigma_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \mu_i)(v - \mu_j) f_{ij}(y, v) dy\, dv$$

if y_i and y_j are continuous random variables with joint density function $f_{ij}(y, v)$ and

$$\sigma_{ij} = \sum_{\substack{\text{all} \\ y}} \sum_{\substack{\text{all} \\ v}} (y - \mu_i)(v - \mu_j) P_{ij}(y, v)$$

if y_i and y_j are discrete random variables with joint probability function

$$p_{ij}(y, v) \;=\; Pr(y_i = y, V_j = v)$$

Result 15.

Let $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ be a random vector with

$$\boldsymbol{\mu} = E(\mathbf{y}) \quad \text{and} \quad \boldsymbol{\Sigma} = Var(\mathbf{y}),$$

and let

$$\mathbf{A}_{p\times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pn} \end{bmatrix}$$

be a matrix of non-random elements, and let

$$\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \quad \text{and} \quad \mathbf{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_p \end{bmatrix}$$

be vectors of non-random elements, then

(i) $E(\mathbf{Ay}+\mathbf{d}) = \mathbf{A}\boldsymbol{\mu}+\mathbf{d}$

(ii) $Var(\mathbf{Ay}+\mathbf{d}) = \mathbf{A\Sigma A}^T$

(iii) $E(\mathbf{c}^T\mathbf{y}) = \mathbf{c}^T\boldsymbol{\mu}$

(iv) $Var(\mathbf{c}^T\mathbf{y}) = \mathbf{c}^T\boldsymbol{\Sigma}\mathbf{c}$

Example 29.

Let the 3×1 random vector $\mathbf{y}$ follows a multivariate normal distribution with men vector $\boldsymbol{\mu} = [7 \quad 9 \quad 5]^T$ and covariance matrix Σ where

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & 3 \end{bmatrix}$$

Consider the vector $\mathbf{w}$ where

$$\mathbf{w} = \begin{bmatrix} 3y_1 - y_2 + 2y_3 - 25 \\ 2y_1 + y_2 - 4y_3 - 12 \end{bmatrix}$$

(a) Find the mean vector of $\mathbf{w}$.

(b) Find the covariance matrix of $\mathbf{w}$.